

## FIGURE 177

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA55737

><subunit 1 of 1, 1141 aa, 1 stop

><MW: 124671, pI: 5.82, NX(S/T): 5

MAGARSRDPWGASGICYLFGSLLVELLFSRAVAFNLDVMGALRKEGEPGSLFGFSVALHRQL  
QPRPQSWLLVGAPQALALPGQQANRTGGLFACPLSLEETDCYRVDIDQGADMQKESKENQWL  
GVSVRSQGPGGKIVTCAHRYEARQRVDQILETRDMIGRCFVLSQDLAIRDELDDGGEWKFC  
RPQGHEQFGFCQQGTAAAFSPDSHYLLFGAPGTYNWKGATARVELCAQGSADLAHLDDGPYEA  
GGEKEQDPRLI PVPANSYFGFSIDSGKGLVRAEELS FVAGAPRANHKGAVVILRKDSASRLV  
PEVMLSGERLTSGFGYSLAVADLNSDGWPD LIVGAPYFFERQEELGGAVVYVLNQGHWAGI  
SPLRLCGSPDSMFGISLAVLGD LNQDGF PDIAVGAPFDGDGKVFIYHGSSLGVVAKPSQVLE  
GEAVGIKSFGYSLSGSLDMDGNQYPDLLVGLADTAVLFRARPILHVSHEVSIAPRSIDLEQ  
PNCAGGHSVCVDLRVCFSYIAVPSSYSPTVALDYVLDADTDRLRGQVPRVTFLSRNLEEPK  
HQASGTVWLKHQHDRVCGDAMFQLQENVKDKLRAIVVTLSYSLQTPRLRRQAPGQGLPPVAP  
ILNAHQPS TQRAEIHFLKQCGCEDKICQSNLQLVHARFCTRVSDTEFQPLPMDVDGTTALFA  
LSGQPVIGLELMVTNLPSDPAQPQADGDDAHEAQLLVMLPDSLHYSQVRLDPAEKPLCLSN  
ENASHVECELGNPMKRGAQVTFYLILSTSGISIIETTELEVELLATISEQELHPVSARARVF  
IELPLSIAGMAIPQQLFFSGVVRGERAMQSERDVGSKVKYEVTVSNOGQSLRTLGS AFLNIM  
WPHEIANGKWL LYPMQVELEGGQGPQKGLCSRPNIHLHDVDSRDRRRRELEPPEQOEPEGE  
RQEPSMSWWPVSSAEKKKNITLDCARGTANCVVFSCPLYSFDRAAVLHVWGRLWNSTFLEEY  
SAVKSLEIVIRANITVKSSIKNLMLRDASTVIPVMVYLDPMAVVAEGVPWWVILLAVLAGLL  
VLALLVLLWKMGGFFKRAKHPEATVPQYHAVKIPREDRQQFKEEKTGTILRNNWGS PRREGP  
DAHPILAADGHP ELGPDGHPGPGTA

### Important features:

#### Signal peptide:

amino acids 1-33

#### Transmembrane domain:

amino acids 1040-1062

#### N-glycosylation sites.

amino acids 86-89, 746-749, 949-952, 985-988 and 1005-1008

#### Integrins alpha chain proteins.

amino acids 1064-1071, 384-408, 1041-1071, 317-346, 443-465, 385-407, 215-224, 634-647, 85-99, 322-346, 470-479, 442-466, 379-408 and 1031-1047

## FIGURE 178

CGCGCCGGGCGCAGGGAGCTGAGTGGACGGCTCGAGACGGCGGCGCGTGCAGCAGCTCCAGA  
AAGCAGCGAGTTGGCAGAGCAGGGCTGCATTTCCAGCAGGAGCTGCGAGCACAGTGCTGGCT  
CACAACAAGATGCTCAAGGTGTGAGCCGTACTGTGTGTGTGTGCAGCCGCTTGGTGCAGTCA  
GTCTCTCGCAGCTGCCGCGGCGGTGGCTGCAGCCGGGGGGCGGTCCGACGGCGGTAATTTTC  
TGGATGATAAACAATGGCTCACCACAATCTCTCAGTATGACAAGGAAGTCGGACAGTGGAAC  
AAATTCCGAGACGAAGTAGAGGATGATTATTTCCGCACTTGGAGTCCAGGAAAACCTTCGA  
TCAGGCTTTAGATCCAGCTAAGGATCCATGCTTAAAGATGAAATGTAGTCGCCATAAAGTAT  
GCATTGCTCAAGATTCTCAGACTGCAGTCTGCATTAGTCACCGGAGGCTTACACACAGGATG  
AAAGAAGCAGGAGTAGACCATAGGCAGTGGAGGGGTCCCATATTATCCACCTGCAAGCAGTG  
CCCAGTGGTCTATCCCAGCCCTGTTTGTGGTTCAGATGGTCATACCTACTCTTTTCAGTGCA  
AACTAGAATATCAGGCATGTGTCTTAGGAAAACAGATCTCAGTCAAATGTGAAGGACATTGC  
CCATGTCCTTCAGATAAGCCCACCAGTACAAGCAGAAATGTTAAGAGAGCATGCAGTGACCT  
GGAGTTCAGGGAAGTGGCAAACAGATTGCGGGACTGGTTCAAGGCCCTTCATGAAAGTGGAA  
GTCAAAACAAGAAGACAAAAACATTGCTGAGGCCTGAGAGAAGCAGATTCGATACCAGCATC  
TTGCCAATTTGCAAGGACTCACTTGGCTGGATGTTTAAACAGACTTGATACAACTATGACCT  
GCTATTGGACCAGTCAGAGCTCAGAAGCATTTACCTTGATAAGAATGAACAGTGTACCAAGG  
CATTCTTCAATTCTTGTGACACATACAAGGACAGTTTAATATCTAATAATGAGTGGTGCTAC  
TGCTTCCAGAGACAGCAAGACCCACCTTGCCAGACTGAGCTCAGCAATATTCAGAAGCGGCA  
AGGGGTAAAGAAGCTCCTAGGACAGTATATCCCCCTGTGTGATGAAGATGGTTACTACAAGC  
CAACACAATGTCATGGCAGTGTGAGCAGTGCTGGTGTGTTGACAGATATGGAAATGAAGTC  
ATGGGATCCAGAATAAATGGTGTGTCAGATTGTGCTATAGATTTTGAGATCTCCGGAGATTT  
TGCTAGTGGCGATTTTCATGAATGGACTGATGATGAGGATGATGAAGACGATATTATGAATG  
ATGAAGATGAAATTGAAGATGATGATGAAGATGAAGGGGATGATGATGATGGTGGTGATGAC  
CATGATGTATACATTTGATTGATGACAGTTGAAATCAATAAATTTCTACATTTCTAATATTTA  
CAAAAATGATAGCCTATTTAAATTTATCTTCTTCCCCAATAACAAAATGATTCTAAACCTCA  
CATATATTTTGTATAATTATTTGAAAAATTGCAGCTAAAGTTATAGAACTTTATGTTTAAAT  
AAGAATCATTTGCTTTGAGTTTTTATATTCTTACACAAAAGAAAATACATATGCAGTCTA  
GTCAGACAAAATAAAGTTTTGAAGTGCTACTATAATAAATTTTTTACGAGAACAACTTTGT  
AAATCTTCCATAAGCAAAATGACAGCTAGTGCTTGGGATCGTACATGTTAATTTTTTTGAAAG  
ATAATTCTAAGTGAAATTTAAATAAATAAATTTTTTAATGACCTGGGTCTTAAGGATTTAGG  
AAAAATATGCATGCTTTAATTGCATTTCCAAAGTAGCATCTTGCTAGACCTAGATGAGTCAG  
GATAACAGAGAGATACCACATGACTCCAAAAAAAAAAAAAAAAA

[illegible]

```
><subunit 1 of 1, 436 aa, 1 stop
```

MLKVSAVLCVCAAAWCSQSLAAAAVAAGGRSDGGNFLDDKQWLTTISQYDKEVGQWNKFR  
DEVEDDYFRTWSPGKPFQALDPAKDPCLKMKCSRHKVCIAQDSQTAVCISHRRLTHRMKEA  
GVDHRQWRGPILSTCKQCPCVVYPSVPCGSDGHTYSFQCKLEYQACVLGKQISVKCEGHCP  
SDKPTSTRNVKRACSDLEFREVANRLRDWFKALHESGSQNKKTLLRPERSRFDTSILPI  
CKDSLGMFNRDLDTNYDLLLDQSELRSIYLDKNEQCTKAFFNSCDTYKDSLISNNEWCYCFQ  
RQQDPPCQTELSNIQKRQGVKLLGQYIPLCEDEDGYKYPTQCHGSVGCWCVDRYGNEVMGS  
RINGVADCAIDFEISGDFASGDFHEWTTDDEDDDDIMNDEDEIEDDDDEDEGDDDDGGDDHDVYI

**Signal peptide:**

Leucine zipper pattern.

**N-myristoylation sites.**

amino acids 357-362, 371-376 and 376-381

## Thyroglobulin type-1 repeat proteins

amino acids 353-365 and 339-352